

UNITED STATES PATENT APPLICATION

for

APPARATUS AND METHOD FOR VOICE MESSAGE CONTROL

Inventors:

Tim A. Williams
Daniel A. Brookshire

prepared by:

BLAKELY, SOKOLOFF, TAYLOR & ZAFMAN LLP
12400 Wilshire Boulevard
Los Angeles, CA 90025-1026
(408) 720-8300

File No.: 005887.P001

EXPRESS MAIL CERTIFICATE OF MAILING

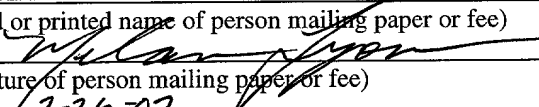
"Express Mail" mailing label number: EL 867 648 085 US

Date of Deposit: February 26, 2002

I hereby certify that I am causing this paper or fee to be deposited with the United States Postal Service "Express Mail Post Office to Addressee" service on the date indicated above and that this paper or fee has been addressed to the Assistant Commissioner for Patents, Washington, D. C. 20231

MELANIE LYONS

(Typed or printed name of person mailing paper or fee)


(Signature of person mailing paper or fee)

2-26-02
(Date signed)

APPARATUS AND METHOD FOR VOICE MESSAGE CONTROL

FIELD OF THE INVENTION

[0001] The present invention relates to the field of communications; more particularly, the present invention relates to performing communications with a mobile device using a packet data channel.

BACKGROUND OF THE INVENTION

[0002] There are a number of technologies available for transferring text and voice information. For example, to transfer text information in real time, NetMeeting from Microsoft of Redmond, Washington may be used. Similarly, if non-real time text transfer is desired, but relatively quick communication in the approximate one to fifteen minute time frame is desired, then AOL Instant Messenger (AIM), Short Messaging Service over Cellular Networks (SMS) or paging (e.g., two-way paging, one-way paging) may be used.

[0003] If a longer period of delay is allowable, text information may be transferred using electronic mail (email) systems. Email systems always have to store a message and then have a recipient retrieve the message to access it. Also, there is no way to know if an email message from a specific person has been received until the email messages are retrieved. One email system disclosed in (Etrieve cite to be added) describes attaching a voice file to an email. The user receives notification of the email by a SMS messaging system, and when the email is responded to, the system retrieves the voice file from memory and plays back the voice file over a circuit switch voice

channel. Therefore, even in this email system, it is still required in this system that the message (the voice file) requires the user to actively retrieve the voice file from a storage facility.

[0004] Long term archival of text messages is a common occurrence and may be performed by using, for example, CD-ROM. Long term archival of voice messages, however, is not performed today with the capability to effectively index the messages.

[0005] Many systems exist for transferring voice information. For example, in real-time voice transfer, a phone, wired or wireless, may be used. One of the wireless cellular carrier networks, Nextel, currently markets a cellular phone based system that includes two-way radio functionality that permits the user, by pressing a button, to use the phone as a two-way radio to transfer voice to preassigned individuals. Similarly, with respect to voice, there are a number of store and retrieve options for transferring voice such as, for example, voice mail. Also, with respect to archiving, there are a number of ways, such as CD-ROMs and tapes, that may be used to record voice files for archival purposes. However, with respect to the communication window of one to fifteen minutes, there seems to be no counterpoint in voice transfer technology that matches or equates to that of instant messaging, SMS or paging used in the transfer of text messages.

SUMMARY OF THE INVENTION

[0006] A method and apparatus for performing voice message control is described. In one embodiment, the method comprises activating a device, capturing a voice message using a microphone on the device in response to activation of the device, storing at least a portion of the voice message in a memory on the device in response to activation of the device, packetizing the voice message on the device in response to activation of the device, and sending at least one packet of the packetized voice message to a network carrier wirelessly over a packet data network channel in response to activation of the device.

[0007] In an alternative embodiment, the method comprises de-packetizing a packet stream comprising a voice message recorded and packetized by a device and determining an action to take with respect to the voice message based on content in the voice message.

BRIEF DESCRIPTION OF THE DRAWINGS

[0008] The present invention will be understood more fully from the detailed description given below and from the accompanying drawings of various embodiments of the invention, which, however, should not be taken to limit the invention to the specific embodiments, but are for explanation and understanding only.

[0009] **Figure 1** illustrates an exemplary architecture of a communication system.

[0010] **Figure 2** is a flow diagram of one embodiment of a process performed by a mobile device (or other device with communication capabilities) in a network environment.

[0011] **Figure 3** is one embodiment of a mobile device.

[0012] **Figure 4** is a flow diagram of one embodiment of a process performed by a mobile device to process menu items.

[0013] **Figure 5** is a flow diagram of one embodiment of a process for routing a voice message.

[0014] **Figure 6** is a flow diagram of one embodiment of the process to identify an operation and specified recipient(s).

DETAILED DESCRIPTION OF THE PRESENT INVENTION

[0015] A communication system is described in which a user of a mobile device, such as a cellular phone, to put the phone in a particular mode, such as by pressing a button on the phone, and causing an audio (voice) message to be queued, sent over a packet data network channel and routed to a recipient or location specified in the message according to a pre-specified routing mechanism. The routing mechanism may cause the message to be forwarded to, for example, another cellular phone in the same carrier network, pager or other mobile device in a different carrier network, a telephone that is part of a Plain Old Telephone System (POTS), a personal digital assistant (PDA), a VoP terminal, or any voice capable device communicating via wireless LAN technologies.

[0016] In the following description, numerous details are set forth to provide a thorough understanding of the present invention. It will be apparent, however, to one skilled in the art, that the present invention may be practiced without these specific details. In other instances, well-known structures and devices are shown in block diagram form, rather than in detail, in order to avoid obscuring the present invention.

[0017] Some portions of the detailed descriptions that follow are presented in terms of algorithms and symbolic representations of operations on data bits within a computer memory. These algorithmic descriptions and representations are the means used by those skilled in the data processing arts to most effectively convey the substance of their work to others skilled in the art. An algorithm is here, and generally, conceived to be a self-consistent sequence of steps leading to a desired result. The steps

are those requiring physical manipulations of physical quantities. Usually, though not necessarily, these quantities take the form of electrical or magnetic signals capable of being stored, transferred, combined, compared, and otherwise manipulated. It has proven convenient at times, principally for reasons of common usage, to refer to these signals as bits, values, elements, symbols, characters, terms, numbers, or the like.

[0018] It should be borne in mind, however, that all of these and similar terms are to be associated with the appropriate physical quantities and are merely convenient labels applied to these quantities. Unless specifically stated otherwise as apparent from the following discussion, it is appreciated that throughout the description, discussions utilizing terms such as "processing" or "computing" or "calculating" or "determining" or "displaying" or the like, refer to the action and processes of a computer system, or similar electronic computing device, that manipulates and transforms data represented as physical (electronic) quantities within the computer system's registers and memories into other data similarly represented as physical quantities within the computer system memories or registers or other such information storage, transmission or display devices.

[0019] The present invention also relates to apparatus for performing the operations herein. This apparatus may be specially constructed for the required purposes, or it may comprise a general purpose computer selectively activated or reconfigured by a computer program stored in the computer. Such a computer program may be stored in a computer readable storage medium, such as, but is not limited to, any type of disk including floppy disks, optical disks, CD-ROMs, and

magnetic-optical disks, read-only memories (ROMs), random access memories (RAMs), EPROMs, EEPROMs, magnetic or optical cards, or any type of media suitable for storing electronic instructions, and each coupled to a computer system bus.

[0020] The algorithms and displays presented herein are not inherently related to any particular computer or other apparatus. Various general purpose systems may be used with programs in accordance with the teachings herein, or it may prove convenient to construct more specialized apparatus to perform the required method steps. The required structure for a variety of these systems will appear from the description below. In addition, the present invention is not described with reference to any particular programming language. It will be appreciated that a variety of programming languages may be used to implement the teachings of the invention as described herein.

[0021] A machine-readable medium includes any mechanism for storing or transmitting information in a form readable by a machine (e.g., a computer). For example, a machine-readable medium includes read only memory ("ROM"); random access memory ("RAM"); magnetic disk storage media; optical storage media; flash memory devices; electrical, optical, acoustical or other form of propagated signals (e.g., carrier waves, infrared signals, digital signals, etc.); etc.

An Exemplary Architecture

[0022] Figure 1 illustrates an exemplary architecture of a communication system. Referring to Figure 1, the voice messaging communication system may include a mobile

device 101 (e.g., mobile handset, phone, computer, personal digital assistant (PDA), etc.) that is communicably coupled to a wireless carrier's network 103 via circuit switched voice, messaging and packet data network channels 102. In one embodiment, the circuit switched voice channel is a channel which primarily carries digitized and compressed voice represented as bits of information placed into a regular time slot on the channel (a wireline telephony example of a similar structure is that of a single voice channel, DS0, within a the T1 or DS1 carrier, a cellular phone network example is the voice channel of a GSM phone), the messaging channel is used to primarily provide a call setup and roaming function for controlling the operation of mobile device 101, and the packet data network channel is a channel which provides packet data communications capability. In one embodiment, this packet data communications capability has a data rate of between 115 kb/s and 2 Mb/s. In one embodiment, the packet data channel is also used to communicate control information. In such a case, the packet data network channel operates as a digital channel. Alternatively, TDM channels may be transferred as well.

[0023] Carrier network 103 is coupled to the network interface (e.g., the VPN) 107 to Internet (or other network environment) 104. In one embodiment, carrier network 103 is WAP-enabled to allow Internet connectivity of a mobile device. In this way, WAP and packet data channels can co-exist. A download server 180 may be coupled to carrier network 103. Download server 180 may be used to download software to mobile device 101. This software may comprise a Java 2 Mobile Execution (J2ME)

program or other programs that mobile device 101 may use to process the voice messages and transmit them onto the packet data network channel.

[0024] Messaging server 105 is coupled to network environment 104 via network interface 108. One or more additional network carrier networks, such as carrier networks 120 and 121, providing access to mobile devices 122 and pager 123, respectively, are also communicably coupled to messaging server 105. Messaging server 105 may be communicably coupled to carrier networks 120 and 121 through network environment 104. This may be by Voice Over Packet communications (VOP). A version of VOP communications is known as VoIP. Such communications may be used for communication between messaging server 105 and carrier network 103 as well. In an alternative embodiment, messaging server 105 and one or more of carrier networks 120 and 121 may be co-located. In such a case, communication may occur directly between the parties, as opposed to going through network environment 104.

[0025] One or more connectivity servers 110_1 - 110_N may be coupled to network environment 104. Messaging server 105 communicates with each of connectivity servers 110_1 - 110_N through network environment 104. This communication may be by VOP. In one embodiment, each connectivity server 110_1 - 110_N is coupled to an exchange server (e.g., Microsoft Exchange Server) and also is coupled to storage 112, which may include one or more databases, including a routing database and an archival database. These databases may be stored in the same memory or separate memories.

[0026] Each connectivity server 110_1 - 110_N may be coupled to a PBX, such as PBX 111, which may include a voice mail system, to provide access to telephones within the

PBX as well as circuit switched access to the PSTN or packet based access to other voice services, such as telephone 140. Note that some embodiments of connectivity servers may or may not include all the features shown in Figure 1 and described herein.

[0027] Connectivity server 110₁ is shown having access to an instant messaging unit 150 to use instant messaging, a wireless local area network (LAN) to communicate with a device accessible thereby, and a workstation 152 to contact PDA 153.

[0028] A point of presence (POP) 133 is also coupled to network environment 104 to provide access via Voice Over Packet (VOP) to telephones, such as telephone 140.

[0029] A voice file archive 132 is also coupled to the network environment 104 to archive voice messages. In one embodiment, communication between messaging server 105 and voice file archival is by VOP.

[0030] Messaging server 105 is coupled to SMS functional unit 154 and instant messaging functional unit 155, which provide access to SMS and instant messaging capabilities, respectively, to messaging server 105.

[0031] Messaging server 105 is also coupled to speech recognition processor 106, and optionally coupled to computer system 131, routing database 117, and an archival database 118. Computer system 131 may be coupled to messaging server 105 directly or through one or more intermediaries, i.e., such as network environment 104 (via, for example, web access) to set up routing information for individuals to be stored in routing database 117 or to access and manage (e.g., delete) voice messages stored in archival database 118.

[0032] Note that the term "server" as used herein is not limited to a single computer system executing software and may comprise one or more software processes running on one or more different computer systems.

[0033] In one embodiment, routing database 117 stores a routing address book of routing information specifying the communication mechanism that is to be used by messaging server 105 to forward a voice message during specific times of each day, week, month and/or year. For example, for one individual, the routing information may indicate that from 8:00 – 10:00 a.m. all voice messages should be forwarded to their regular land-line telephone via a wired line, (e.g., telephone 140 via PBX 111 accessed through corporate server 110 or POP 133), from 10:00 a.m. to 5:00 p.m. all voice messages should be forwarded to their cell phone via a specified carrier network (e.g., mobile 122 via carrier network 120), from 5:00 p.m. – 7:00 p.m. all voice messages should be forwarded to their pager via a specified carrier network (e.g., pager 123 via carrier network 121), and from 7:00 p.m. to 8:00 a.m. all voice messages should be forwarded to archival database 118 (or voice message archive 132) for storage as a voice mail message for later retrieval. This routing information may be part of each user profile maintained in the system.

[0034] In one embodiment, the communication architecture described in Figure 1 enables the user of a mobile device, such as mobile device 101, to perform one or more of the following types of communications: 1) an interpersonal communication (send to another person); 2) a group communication (send to a group of people, such as an engineering work group); and 3) memo to self; 4) interactions with computers.

Examples of interaction with computers include access to scheduling and calendaring information that may be contained within a user's Outlook (e.g., Microsoft Outlook) program on the user's desktop computer or within the user's PDA. Another example of interaction with computers is allowing access to the user's account on a voice mail system for the purposes of control, message retrieval, and/or message storage.

Interpersonal Communications

[0035] To perform an interpersonal communication to communicate with another individual in a store and forward manner, a user of mobile device 101 activates mobile device 101. Activating mobile device 101 may comprise pressing a button (e.g., key on a keypad, soft button (e.g., touch screen touched by a finger)) or using some other selection mechanism (e.g., stylus, mouse click, speech recognition on the handset, etc.) on mobile device 101. Activating mobile device 101 may comprise receiving an authorization from a biometric device (e.g., a speech recognizer to identify an individual by their voice).

[0036] In response to this activation (e.g., selection), mobile device 101 causes utterances (a voice message) to be queued and sent as a voice file from mobile device 101 via a packet data channel and forwarded to another individual. In response to the button being pressed on mobile device 101, a voice message may be created and sent over packet data channel 102 to carrier network 103. Thus, pressing the button on mobile device 101 activates the packet data channel without dialing a phone number

and mobile device 101 is able to send a voice message to another without having to perform any phone number lookup.

[0037] Carrier network 103 separates the packets received from mobile device 102 and sends them to messaging server 105. In one embodiment, a firewall of carrier network 103 normally allows unimpeded access to Internet 104. In one embodiment, carrier network 103 uses a virtual private network (VPN) connection (i.e., a port on the firewall of carrier network 103) to Internet 104 to send the packetized voice message received over the data packet network channel from mobile device 101. Carrier network 103 may perform a network address translation (NAT) to identify a packet stream from mobile device 101 as one to be forwarded to Internet 104.

[0038] Messaging server 105 determines actions to take with the voice message based on its contents. For example, a user of mobile device 101 may record a voice message such as "Call Mary engineering meeting is canceled." In response to receiving this message, messaging server 105 determines that a call is to be made to a specified recipient named Mary. In order to complete this call, messaging server 105 is able to determine who the specified recipient(s) (e.g., Mary in this example) is and how to contact the specified recipient(s).

[0039] Messaging server 105 may use speech recognition on the voice message to identify names of individuals contained in the message as well as one or more commands. In one embodiment, messaging server 105 knows the portion of the voice messages that are command words (or phrases) and names of specified recipients by constraining the command words (or phrases) to a predetermined set and constraining

the location in the voice message of both the command words and named recipients (or entity). More specifically, constraining the context of the sentence, for example the first word is always one of a small set of words (e.g., call, schedule, forward, memo) followed by the recipient name as it is contained within the routing address book. The commands are identified by comparing recognized words with a list of preselected command words and individual words are parsed by the intervening silence.

[0040] In one embodiment, if the first word is not one of the predefined set of words, messaging server 105 saves the voice message and sends a menu list to the user of what actions are to be taken, e.g., call, schedule, forward, memo, and a list of recipients from the address book, if that is necessary. In another embodiment, if the speech recognizer cannot adequately determine the contents of the voice message, the voice message is routed to a human operator who performs the speech-to-text processing by listening to the message and transcribing it into text. The voice message may have digital signal processing performed on it prior to being routed to the human operator. An example of which is the reduction of background noise. Thus, messaging server 105 may reflect back to mobile device 101 a textual list of commands and/or recipients in response to the voice message if it was not clear after performing speech recognition who the specified recipient(s) is or the command(s) that is to be performed as a prompt to the user to clarify the desired command and/or recipient(s), if any. In such a case, messaging server 105 generates a text message with a command recognizable to the mobile device and sends the text message to carrier network 103, which forwards the message to mobile device 101. The text message may be sent to

mobile device 101 over the messaging or packet channel. In one embodiment, the prompt can come either through WAP (packet channel), which causes the prompt to be presented on a static web page like browser interface, or in alternative embodiments, it can come through the packet channel to a JAVA or other similar program running on mobile device 101 that displays the prompt (e.g., menu) on a display of mobile device 101.

[0041] Messaging server 105 determines how to route the voice message to the specified recipient(s) by locating routing information for the specified recipient(s). In one embodiment, messaging server 105 accesses a local database, such as routing database 117, using the name of the specified recipient(s), to obtain the necessary routing information from a previously entered profile as specified by the user.

[0042] In an alternative embodiment, messaging server 105 locates the routing information for the specified recipient(s) by contacting one of the corporate servers. The corporate server maintains routing information for a number of individuals in a database. Messaging server 105 sends the name(s) of the specified recipient(s) and the sender to the corporate server, which accesses its database and provides the requested routing information. In one embodiment, the corporate server may use Microsoft Exchange Server or other similar functioning server to identify the routing information for the specified recipient(s) in response to receiving the name(s) of the specified recipient.

[0043] Note that if more than one corporate server is maintaining routing information, messaging server 105 identifies the corporate server that is storing the

routing information for the specified recipient(s) it needs based on a unique identifier associated with the mobile device sending the voice message which identifies the user who is originating the message. More specifically, in one embodiment, each user is assigned a unique identifier and this unique identifier is included in the packet header of the packets containing the voice message that is sent on the packet data network channel. When messaging server 105 receives the packets, it obtains this unique identifier and accesses a local memory that is able to associate a corporate server with the unique identifier. In one embodiment, the local memory includes a listing of all unique identifiers and their associated corporate server. In an alternative embodiment, a hash table is used and the unique identifier is used to hash to a value indicative of the corporate server associated with that unique identifier.

[0044] Thus, messaging server 105 determines how to route the voice file message to the recipient(s) specified in the voice message and routes the voice file to the specified recipient(s). Thus, the voice messages route themselves in that the information needed to determine where to route the messages is determined using the content of the voice message. For example, the determination of how to route the voice file to Mary may be based on local information, such as the information stored in the routing database 117, to which messaging server 105 has access, or may be determined by accessing another server, such as one of connectivity servers 110₁-110_N. In the latter case, messaging server 105 would forward the name Mary to the corporate server, which would access a routing database, such as a routing database in storage 112₁ and

provide information indicative of how to route a message to Mary back to messaging server 105. Using that information, messaging server 105 routes the message to Mary.

[0045] The routing information may indicate that any voice message is to be routed to the specified recipient by way of another mobile device accessible via carrier network 103. In such a case, upon determining the specified recipient and the routing information specifying a mobile device in the coverage area of the carrier network 103, messaging server 105 sends a packetized stream through carrier network 103 via network environment 104, to be sent to the mobile device.

[0046] In one embodiment, messaging server 105 contacts the mobile device using the circuit switched channel in a typical fashion, such as by calling the mobile device. When the individual answers, messaging server 105 plays a voice prompt telling the individual that a voice message exists for the individual and asks whether the individual will like to hear the voice message. The individual may be instructed to indicate their desire to hear the message in one or more ways, such as, for example, by pressing a particular button on the mobile device, saying a particular phrase (which would be recognizable by messaging server 105), or selecting a menu item displayed on the phone. In response to the selection, messaging server 105 plays the message.

[0047] In an alternative embodiment, the packetized stream is sent to the mobile device via through carrier network 103 using the packet data network channel. In such a case, the mobile device includes functionality to play or review the voice message if sent via the packet data network channel. Such functionality includes a de-packetizer to depacketize the stream to retrieve the voice message and an audio player to operate in

conjunction with any speaker of the mobile device to generate audio signals to drive the speaker to play the voice message.

[0048] In one embodiment, voice mail-like controls of play, skip, fast forward, backup, delete, and reply will be available to the user at the time of reviewing the voice messages regardless of the delivery mechanism of packet channel or circuit switched channel.

[0049] If the routing information indicates that the specified recipient is at a POTS telephone or a PBX station set, such as telephone 140, messaging server 105 may route the voice message to telephone 140 using Voice Over Packet (VOP) to POP 133 and onto telephone 140, or may gain access to a corporate servers' PBX, such as PBX 111, and utilize the connectivity server 110₁ to initiate the call to telephone 140. In either case, messaging server 105 converts the packet data to analog voice to play the voice message.

[0050] If the routing information indicates that the specified recipient is on a mobile device of another carrier network, messaging server 105 may initiate a call to that other mobile device. For example, it specifies individuals at mobile phone 122, messaging server 105 may initiate the call through to carrier network 120 in order to place the call to mobile device 122 in the same way the call is made and the message is delivered as described above. That is, if a packet data network channel is not being used, messaging server 105 may convert the voice message to analog speech using the an appropriate converter and send a call to mobile device 122 using a circuit switch voice channel. Further, alternatively, messaging server 105 may send use a voice-to-text

converter to generate text messages and send it to the mobile device via a messaging or packet channel, if such a messaging or packet channel is available.

[0051] If the specified recipient is on a device such as (one-way or two-way) pager 123, messaging server 105 converts the voice file to text and sends the text as a text message to the pager through its carrier network (e.g., pager 123 through carrier network 121).

[0052] Note that, in one embodiment, if an individual declines to receive a voice message after being prompted regarding its availability or does not respond to the call from messaging server 105, messaging server 105 may store the message into the individual's voice message storage archival facility, such as voice mail archive 132, or has the message played into a voice mail system, such as voice mail 111A by connectivity server 110₁. This connection with the voice mail system 111a is performed by the connectivity server. One method to perform this operation is for the connectivity server to place a phone call (circuit switched or VOP) into the PBX essentially dialing phone number corresponding to the user's voice mail box extension. In one embodiment, when a voice message is archived, the voice message is tagged with the date and time of the voice message, as well as the sender and specified recipient(s) of the voice message and message length and priority.

Group Communications

[0053] Group communications may be performed in the same manner as interpersonal communications except that the specified recipient of the voice message

received by messaging server 105 comprises the name of a group or a multiplicity of recipients. In such a case, in one embodiment, routing server 105 or corporate server 105 includes a database listing created by the sender or surrogate of each individual in the group and obtains the routing information for each of the individuals in the group. Using the routing information for each of the individuals in the group, messaging server 105 forwards the voice message to each individual as individual communications. Thus, if the routing information in each of the specified recipients' profiles are to multiple devices, including different types of devices (e.g., cellular phone, pager, landline telephones, etc.), messaging server 105 routes the message to each device as a separate communication.

[0054] Alternatively, messaging server 105 uses the unique identifier in the packet header to identify a corporate server and sends the group name to the corporate server. In response, the corporate server sends the routing information for each of the members in the group to messaging server 105 so that messaging server 105 is able to route the voice file to the individuals in the group correctly.

Memos

[0055] The architecture may enable an individual to send themselves a memo. In such a case, the user of a mobile device, such as mobile device 101, presses a button or other selection mechanism on their mobile device to record a voice message with an indication that the voice message is a memo. This voice message is then packetized and

sent to messaging server 105, which identifies it as a memo and stores the memo in an archive (e.g., archive 132, archive 118, etc.).

[0056] Memos may be retrieved by the individual in the same way as a voice message or the memo may be scheduled to return to the user at a specific time and date. In one embodiment, a browser interface may be used to access and review messages, including memos. This browser interface allows the user to audio playback the message and/or have it converted to text and displayed.

[0057] Alternatively, individuals may forward memos to other people.

[0058] In one embodiment, messaging server 105 automatically creates an email to the mobile device user by converting the voice file to text and sending the email to the user via normal email facilities.

[0059] If an Outlook-based system is employed, a reminder or notification may be launched automatically from Outlook. This is performed by the connectivity server obtaining information from the user's Calendar or PIM (Personal Information Management) system (e.g., Microsoft Outlook) regarding the onset of a calendar or memo event. The connectivity server associates the event with a voice file and schedules a voice message to be transmitted to the user. The voice file can either be a prerecorded message or be created from the event itself via a text-to-speech system associated with or part of the messaging server.

[0060] Note that in alternative embodiments, the voice messaging described herein may be performed with a device that is not a mobile device. For example, the voice messaging may be performed with a PSTN phone. In such a case, the PSTN

phone dials into messaging server 105 and leaves a message. Messaging server 105 processes the message in the same manner as if received from a mobile device.

Other Features of the Architecture

[0061] In one embodiment, messaging server 105 archives voice messages and other information for billing purposes. Such information may be archived using database 118 or voice message archive 132. Similarly, corporate server 110₁-110_N may include a portion of storage 112₁-112_N, respectively, for use as an archive.

[0062] In one embodiment, download server 112 enables over-the-air download of software modules, such as for example, J2ME, to reconfigure a mobile device. In such a case, download server 112 downloads software to carrier network 1xx, which sends the software to a mobile device, such as mobile device 102. Therefore, even if mobile device 102 is not initially programmed to engage in the non-real time communication described herein, it can be after being deployed. More specifically, in one embodiment, each carrier network includes a specific MIME number for a particular application run by the mobile device. The MIME number allows a user browsing the world wide web on the cell phone to cause an application to be downloaded to the cell phone for use.

Exemplary Flow Processing

[0063] Figure 2 is a flow diagram of one embodiment of a process performed by a mobile device in a network environment. The process is performed by processing logic

which may comprise hardware (e.g., circuitry, dedicated logic, etc.), software (such as run on a general purpose computer system or a dedicated machine), or a combination of both.

[0064] Referring to Figure 2, processing logic in a mobile device receives an activation indication (processing block 201). In one embodiment, such an activation may be received in response to the pressing of a button on the mobile device. The button may comprise a key on a keypad. In response to receiving the activation, the processing logic captures utterances (voice) (processing block 202) and stores the captured utterances in a file as a voice message (processing block 203). Subsequently, processing logic in a mobile device packetizes the voice file (processing block 204) and sends the packet flow to the network carrier (processing block 205).

[0065] Figure 3 is one embodiment of a mobile device, such as mobile device 101. Referring to Figure 3, the user depresses a button or key, performs a stylus selection, or uses some other activation mechanism 309 that signals to controller 307 to operate in a non-real time mode. In response to depression of the button or other activation, microphone 301 records utterances or other audio information and stores the recorded utterances in storage 302.

[0066] The recorded utterances in storage 302 are packetized by packetizer 303 under control of controller 307 and transmitted wirelessly using transmitter 304 and antennae 305 to the carrier network using a packet data network channel (such as shown in Figure 1). Packetizer 303 may be part of a channel modem on the mobile device that is coupled to transmitter 304. In one embodiment, although not shown, a

codec and digital signal processor (DSP) may be included, where the DSP performs LPC coding on the recorded stream of utterances (prior to packetization) in a manner well-known in the art. In an alternative embodiment, the data stream may be processed by a codec and then the digital signal processing may be performed along with the packetization by a process running on processor 306.

[0067] In one embodiment, the recorded utterances stored in storage 302 undergo speech recognition using speech recognition 303. The recognized work are stored back in storage 302 or provided directly to packetizer 303.

[0068] In one embodiment, controller 307 and packetizer 303 are part of the processor 306. More specifically, processor 306 runs software that can set up and launch calls. This software packetizes voice input and causes the packets to be sent on to a data packet channel. Thus, in one embodiment, this software may include the functions performed by controller 307 and packetizer 303. In one embodiment, processor 306 executes a Java 2 Mobile Execution (J2ME) program such that the mobile device functions as a thin client. In one embodiment, the J2ME program (or another program executed by processor 306) includes a speech recognition routine to perform the speech recognition associated with speech recognition 303.

[0069] At times, such as when the messaging server is providing menu options to the user, a mobile device, mobile device utilizes a received path that includes receiver 310 that receives a service of packets from the messaging server that are depacketized using depacketizer 311 and stored in storage 314. Control 307 accesses the packets in storage 314 and displays them on display 312 as a menu selectable by the user. The user

may use selection indication mechanism 313 to make a selection of one of the menu options. In one embodiment, the selection indication mechanism 313 may comprise a cursor control device, a keypad device, stylus, or other well known input device for selecting menu options on a display screen. The result of the selection sent by controller 307 to packetizer 303 and transmitted back out on packet data network channel to the messaging server.

[0070] Although not shown, the coupling of antennae to 305 to transmitter 304 and receiver 310 is usually through a switch or duplexer.

[0071] Figure 4 is a flow diagram of one embodiment of a process performed by a mobile device to process menu items from a messaging server. The process is performed by processing logic which may comprise hardware (e.g., circuitry, dedicated logic, etc.), software (such as run on a general purpose computer system or a dedicated machine) or a combination of both.

[0072] Referring to Figure 4, processing logic in a mobile device receives packets from the messaging server via the packet data network channel (processing block 211). In an alternative embodiment, the information from the messaging server is sent through the network carrier to the mobile device via a messaging or packet channel.

[0073] In response to receiving packets on the packet data network channel, processing logic in the mobile device de-packetizes the packets (processing block 212) and displays the menu with choices based on the information in the packets (processing block 213).

[0074] Subsequently, in response to a user selection, the processing logic in the mobile device receives the selection of a menu item (processing block 214), packetizes the selection (processing block 215), and sends the packets that include the selection to the messaging server via the packet data network channel and the carrier network (processing block 216).

[0075] If the menu is sent on the messaging channel, the user is able to respond by sending a responding message on the message channel in a well-known manner. Assuming the user selects one of the available menu options, the messaging server is able to comprehend the selection based on the fact that the messaging server sent the menu.

Voice Message Routing

[0076] Figure 5 is a flow diagram of one embodiment of a process to route a voice message. The process is performed by processing logic which may comprise hardware (e.g., circuitry, dedicated logic, etc.), software (such as run on a general purpose computer system or a dedicated machine), or a combination of both. The process may be performed by messaging server 105 of Figure 1, which runs software.

[0077] Referring to Figure 5, processing logic in the messaging server depacketizes the packet stream containing a voice file received from mobile device, such as mobile device 101. The depacketizing may be performed by processor, general purpose or dedicated, running a depacketization module (routine). Alternatively, a depacketizer unit may be coupled to messaging server 105.

[0078] Processing logic in the messaging server then performs speech recognition (processing block 502). This may be optional in situations where the voice message received from the mobile device has already undergone speech recognition. The speech recognition may be performed by a speech recognition unit, speech recognition processor running a speech recognition module, or a general purpose processor running a speech recognition module.

[0079] Using the speech recognized information, processing logic in the messaging server may optionally perform parsing to identify key words or phrases in the voice message (processing block 503). Such parsing may be useful in identifying commands or specified recipients associated with the call so that a proper routing of information is performed by the messaging server. The parsing may be performed by a processor, general purpose or dedicated, running a parser module. Alternatively, a parser may be coupled to or associated with the messaging server.

[0080] With the speech recognized voice message, processing logic in the messaging server determines an action to take (processing block 504). In one embodiment, the processing logic determines an action to take by identifying the operation and the specified recipients (processing block 504) and routing the voice message to the specified recipients in the appropriate manner (processing block 504B). The routing may be performed by a processor, general purpose or otherwise, running a communication routing module, in conjunction with communications functionality (e.g., network information cards, transmitters, receivers, etc.) capable of performing all

the necessary communications. Alternatively, the routing may be performed by a communication or routing unit.

[0081] Figure 6 is a flow diagram of one embodiment of the process performed by the messaging server to identify an operation associated with a voice message and one or more specified recipients. The process is performed by processing logic which may comprise hardware (e.g., circuitry, dedicated logic, etc.), software (such as run on a general purpose computer system or a dedicated machine), or a combination of both. In one embodiment, the process is performed by messaging server 105 of Figure 1 running software.

[0082] Referring to Figure 6, the processing logic in the messaging server initially determines whether routing information of the specified recipient(s) is stored locally (processing block 601). If the routing information of the specified recipient(s) is stored locally, processing logic in the messaging server accesses the database using identifiers for the specified individual(s) (processing block 602) and obtains an indication of the manner in which to route the voice message and any necessary information to the specified recipient(s) (processing block 603).

[0083] If the routing information of the specified recipient(s) is not stored locally, processing logic identifies a server (e.g., a connectivity server, a corporate server, etc.) associated with the specified recipient(s) (processing block 611), sends the identifier for the specified person to the identified server (processing block 612), and subsequently receives an indication of the manner in which to route the voice message to the specified recipient(s) and any necessary information to do so (processing block 613).

Switching Between Channels on the Mobile Device

[0084] In one embodiment, when using the mobile device for a circuit switch call, the user may press a button or use another selection mechanism to activate the packet data network channel. In such a case, the circuit switched call is put on hold by the mobile device continuing to process received packets/frames from the circuit switched network while sending idle speech data patterns into the network from the mobile device transmitter. Meanwhile, the speaker and microphone will be utilized by the packet channel process. In one embodiment, the speech decoder/encoder that is coupled between a speaker and a microphone on the mobile device and the mobile device's antenna is left running while its connections between the speaker and microphone are disconnected or disabled. In an alternative embodiment, a signal is sent to the cellular network provider who places the call into the hold state until further notified. When the user is finished with the packet data network channel, then the user presses the button or activates the selection mechanism again and the user is returned to the circuit switched call. This allows for the interruption of a circuit switched call to provide information to the messaging server. Interrupting a call to utilize the packet data channel may be useful, for example, to allow the user to place a caller on hold to make a meeting time notification within his personal information manager (PIM) through the messaging server to the connectivity server to the exchange server and the PIM.

[0085] These communications have a number of characteristics that will be described in more detail below. These characteristics may include, for one or more embodiments:

- 1) the communications are non-real time;
- 2) permit voice and data to the phone;
- 3) support group/chat room interactions;
- 4) may interact with PIM software based (as opposed to typing in the information), which permits a) launching of reminders or notifications from the PIM, b) the scheduling of calendar events (with conflict notification), and c) the ability to access the PIM address book for use in the routing of messages; and
- 5) an instant messaging interface to allow for speech based interaction. This utilizes text to speech and speech to text conversion software.

[0086] Whereas many alterations and modifications of the present invention will no doubt become apparent to a person of ordinary skill in the art after having read the foregoing description, it is to be understood that any particular embodiment shown and described by way of illustration is in no way intended to be considered limiting. Therefore, references to details of various embodiments are not intended to limit the scope of the claims which in themselves recite only those features regarded as essential to the invention.